1    Evaluation of Rainfall-Snowfall Separation Performance in Remote Sensing Datasets

2    Yalei You[1]*, Christa Peters-Lidard[2], Sarah Ringerud[3], John M. Haynes[4]

3

4    [1]Earth System Science Interdisciplinary Center/Cooperative Institute for Climate and Satellites,

5    University of Maryland, College Park, Maryland, USA

6    [2]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

7    [3]Earth System Science Interdisciplinary Center, University of Maryland, College Park,

8    Maryland, USA

9    [4]Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins,

10   Colorado, USA

11

12

13   **Key Points:**

14   • Snowfall determination accuracy varies greatly among four remote sensing datasets

15   ranging from 42% to 96%.

16   • GPM dual frequency radar detects snowfall at about 1.5 km above the ground,

17   instead of surface snowfall.

18   • MERRA2 temperature close to the surface is noticeably colder than observed,

19   leading to more rainfall pixels being classified as snowfall pixels.

20

21

22

23   *Corresponding author: Yalei You, yyou@umd.edu

24    **Abstract**: The first key step to accurately measure global snowfall is to separate rainfall from

25    snowfall correctly (i.e., precipitation phase discrimination). This study first evaluates the phase

26    discrimination performance in four remote sensing datasets, including observations from ground

27    radar, spaceborne radars, and spaceborne radiometer, relative to ground observations. Results

28    show that the snowfall discrimination accuracy varies greatly among these datasets ranging from

29    42% to 96%, dependent on whether and how the temperature information are considered. For

30    example, over half of the snowfall from the GPM spaceborne radar is actually rainfall at the

31    surface since it detects snowfall at ~ 1.5 km above the ground without considering the

32    temperature information close to the surface.  Second, we evaluate the discrimination

33    performance using the temperature information from four reanalysis datasets. It is found that

34    MERRA2 temperature close to the surface is colder than the other three datasets, leading to more

35    rainfall pixels being misclassified as snowfall pixels.

36

37    **Plain Language Summary:** Satellite remote sensing provides the only means of measuring

38    rainfall/snowfall on the global scale.  Misclassifying the precipitation phase (i.e., rainfall as

39    snowfall, or vice versa) could lead to the estimated precipitation rate being one order of

40    magnitude smaller or larger. Our results reveal that the snowfall discrimination accuracy varies

41    greatly among four remote sensing datasets ranging from 42% to 96%. For example, over half of

42    the snowfall from the state-of-the-art precipitation product based on the Global Precipitation

43    Measurement radar is rainfall at the surface without considering the temperature information

44    close to surface. Additionally, the temperature discrepancy among different reanalysis datasets

45    also greatly affects precipitation phase discrimination. Our results show that MERRA2

46    temperature close to the surface is colder than the other three major datasets, leading to more

47    rainfall pixels being misclassified as snowfall pixels.

48

**1. Introduction**

Accurate global snowfall measurement is needed for many applications including water resources management (Gergel et al., 2017), water budget evaluation (Sheffield et al., 2009), and long-term climate change monitoring (OGorman, 2014). Satellite remote sensing provides the only means for the snowfall measurement on the global scale. The CloudSat Cloud Profiling Radar (CPR) observations (Stephens et al., 2002) have been used widely to investigate global snowfall features (Liu, 2008; Kulie and Bennartz, 2009; Kulie et al., 2016, Milani et al., 2018). Snowfall characteristics have also been characterized by the dual frequency precipitation radar (DPR) on board the Global Precipitation Measurement Mission (GPM) Core Observatory Satellite (Adhikari et al., 2018; Skofronick-Jackson et al., 2019), though the GPM DPR has a limited capability to measure light snowfall events with detection sensitivity at about 12 dBZ (Hamada and Takayabu, 2016; Panegrossi et al., 2017; Skofronick-Jackson et al., 2019). In addition, passive microwave radiometers are also commonly used for global snowfall estimation (Kongoli et al., 2003; You et al., 2015; Kummerow et al., 2015; Meng et al., 2017; Kidd et al., 2016; Ebtehaj and Kummerow, 2017). Ground radar observations also provide valuable snowfall information on the regional and continental scale. For example, the Multi-Radar Multi-Sensor (MRMS) provides snowfall estimation over the continental United States at 2-minute and about 1-km resolution (Kirstetter et al., 2012; Zhang et al., 2016).

It remains extremely challenging to accurately measure falling snow from both ground and spaceborne radars (Chen et al., 2016; Skofronick-Jackson et al., 2019; Kulie et al., 2021), and microwave radiometers (Wang et al., 2013; Kummerow et al., 2015; You et al., 2017; Meng et al., 2017; Milani et al., 2021). A major obstacle is how to accurately determine the precipitation

4

72 phase (i.e., separating rainfall from snowfall). Precipitation phase misclassification (e.g., rainfall

73 to snowfall, or vice versa) could lead to the estimated precipitation rate from spaceborne radar

74 being one order of magnitude smaller or larger (Sims and Liu, 2015). On the other hand, there

75 exists very limited information in the remote sensing measurements themselves for precipitation

76 phase determination. Specifically, a bright-band in the radar profiles may indicate rainfall at the

77 surface, while not all radar profiles have a bright-band. Furthermore, the uncontaminated

78 spaceborne radar observations are about 1.5 km above the ground level, while the precipitation

79 phase may change from 1.5 km to the surface. The brightness temperature (TB) observations

80 from passive microwave radiometers do not directly indicate the precipitation phase at the

81 surface. Very cold TB may be associated with hail events (Ferraro et al., 2015; Mroz et al.,

82 2017). However, it is difficult to distinguish snowfall from rainfall by using TB only.

83

84 Researchers have long sought to exploit temperature-related parameters for precipitation phase

85 discrimination. For example, early studies based on limited station observations using the surface

86 air temperature revealed that the rainfall-snowfall transition threshold value is close to 2 $^{\circ}$C

87 (United States Army Corps of Engineers, 1956; Auer Jr, 1974). Later studies showed that other

88 variables besides the surface temperature can also be used to constrain the rainfall-snowfall

89 separation, including surface pressure (Dai, 2008; Sims and Liu, 2015), land surface elevation

90 (Ding et al., 2014; You et al., 2016), relative humidity (Matsuo et al., 1981; Behrangi et al.,

91 2018; Jennings et al., 2018), and temperature lapse rate (Haynes et al., 2009; Sims and Liu,

92 2015). The mean temperature (or thickness) from the surface to the low or mid-troposphere has

93 also been widely used to determine the precipitation type (Wagner, 1957; Bourgouin, 2000). In

94 addition, Jennings et al. (2018) pointed out that the snowfall-rainfall transition temperature varies

95    from -0.4 °C to 2.4 °C across the Northern Hemisphere, further adding complexity and difficulty

96    for accurate precipitation phase partitioning.

97

98    The first objective of this study is to evaluate precipitation phase discrimination performance in

99    four commonly used precipitation remote sensing datasets. They are the CloudSat CPR

100   precipitation product (2C-PRECIP-COLUMN) (Haynes et al., 2009), GPM DPR precipitation

101   product (2ADPR) (Hamada and Takayabu, 2016; Iguchi, 2020; Le et al., 2017), GPM

102   Microwave Imager (GMI) precipitation product (2AGPROFGMI) (Kummerow et al., 2015), and

103   the MRMS precipitation product in the Continental United States (CONUS) (Zhang et al., 2016).

104   More details regarding the precipitation phase determination in these products will be provided

105   in the following sections.  In addition, several phase segregation methods exist in the literatures,

106   and temperature related parameters often are obtained from different reanalysis datasets.

107   Therefore, the second objective is to assess the phase discrimination accuracy by using different

108   methods and reanalysis datasets.

109

110   **2 Datasets and Methodology**

111   2.1 Datasets

112   The reference ("truth") dataset for this study is the precipitation phase code reports (i.e., rainfall

113   or snowfall) from the NOAA National Centers for Environmental Information (NCEI)'s

114   Integrated Surface Database (ISD), which consists of global hourly and synoptic observations

115   from over 35,000 stations worldwide (Smith et al., 2011). This dataset has also been used to train

116   precipitation phase discrimination schemes (Dai, 2008; Liu, 2008; Sims and Liu, 2015; You et

117   al., 2016; Behrangi et al., 2018) and compute precipitation frequencies (Petty, 1995; Dai, 2001).

118

119   The precipitation phase information in the remote sensing datasets to be compared in this study

120   is from CloudSat CPR precipitation product (2C-PRECIP-COLUMN, version 5) (Haynes et al.,

121   2009), GPM DPR precipitation product (2ADPR, version 6) (Hamada and Takayabu, 2016; Le et

122   al., 2017; Iguchi, 2020), GPM Microwave Imager (GMI) precipitation product (2AGPROFGMI,

123   version 5) (Kummerow et al., 2015), and the MRMS precipitation (operational version) (Zhang

124   et al., 2016).

125

126   Specifically, for CPR 2C-PRECIP-COLUMN, we obtain the "Precip_flag" variable, which

127   groups the precipitation phase into one of three categories: rain, snow, and mixed. For DPR, we

128   obtain the "phaseNearSurface" and "flagSurfaceSnowfall". The first variable is determined by

129   the temperature at the near surface level (~1.5 km) and the bright-band information (Iguchi,

130   2020). The second variable is computed from the difference between Ku-band PR (KuPR) and

131   Ka-band PR (KaPR) at the clutter-free height (~1.5 km) (Le et al., 2017). For GPM Microwave

132   Imager (GMI) retrieval from the Goddard Profiling Algorithm (GPROF), we obtain surface

133   precipitation rate ("surfacePrecipitation") and snowfall rate ("frozenPrecipitation"). For MRMS,

134   we obtain the precipitation flag variable ("pcp_flag").

135

136   To collocate the remote sensing datasets with the surface reference data, we use the threshold

137   values of 10 minutes and 10 km for CPR, DPR, and GMI observations. That is, when remote

138   sensing observations and the surface reference data are less than 10 minutes apart in time and

139   less than 10 km away in distance, they are considered as coincident observations. These two

140   threshold values (10 minutes and 10 km) are selected by considering the trade-off between the

141    sample size (primarily for the CPR nadir-only observations) and the accuracy of coincident

142    observations. It is worth mentioning that the spatial resolution for CPR, DPR, and GMI

143    precipitation products are about 1.6 km, 5.2 km, and 14.3 km, respectively. The ancillary

144    temperature information used in these products for precipitation phase determination is often

145    much coarser than these spatial resolutions. For MRMS, we simply find the nearest time and the

146    closest distance with the surface data due to its high spatial (about 1 km) and temporal (2

147    minutes) resolutions. For the collocation period, we use the full CPR observation record from

148    2006 to 2017. For DPR and GMI products, we use observations from March 2014 (launch of the

149    GPM satellite) to December 2018. For MRMS, we only use the data in 2016 with large enough

150    sample size (> 1 million collocated samples) since the ground radars used to generate the MRMS

151    precipitation dataset often are on the same location with the ground gauges, from where the

152    precipitation phase reports are obtained.

153

154    As mentioned previously, the temperature information for the phase determination in remote

155    sensing products is often obtained from model outputs. Our second objective is to evaluate the

156    phase discrimination accuracy when the ancillary temperature information is obtained from

157    different sources. To this end, this study compares four global reanalysis datasets, including the

158    Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2)

159    (Gelaro et al., 2017), the ECMWF Reanalysis 5th generation (ERA5) (Hersbach et al., 2020), the

160    Japanese 55-year Reanalysis (JRA55) (Kobayashi et al., 2015), and the Global Forecast System

161    (GFS) (Kanamitsu et al., 1991). The spatial resolutions are $0.5^{\circ} \times 0.625^{\circ}$ for MERRA-2,

162    $0.25^{\circ} \times 0.25^{\circ}$ for ERA5, $1.25^{\circ} \times 1.25^{\circ}$ for JRA55, and $0.5^{\circ} \times 0.5^{\circ}$ for GFS, respectively. The temporal

163    resolutions for the surface temperature, 2-m air temperature, and the surface pressure are hourly

8

164  for MERRA2 and ERA5, and 6-hourly for JRA55 and GFS, respectively. The temporal

165  resolutions for the temperature profile and geopotential height profile are hourly for ERA5, 3-

166  hourly for MERRA2, and 6-hourly for JRA55 and GFS, respectively. For all these model

167  datasets, we match the surface weather report data with the closest grid and linearly interpolate

168  the temperature information in the temporal dimension.

169

170  **2.2 Evaluation metric**

171  We assess the phase discrimination performance in the four remote sensing datasets using the

172  snowfall or rainfall success percentage. We take the snowfall and the CPR as an example to

173  show the definition of the success percentage. The snowfall success percentage is computed as

174  the number of snowfall observations from CPR divided by the number of the snowfall

175  observations from the reference. Similar computations are applied to rainfall success percentage

176  and to other remote sensing datasets.

177

178  We would like to emphasize that our analysis is not a detection performance evaluation. Instead,

179  the snowfall or rainfall success percentage is computed when both the remote sensing

180  instruments and the surface reference dataset detect precipitation. This requirement normalizes

181  for any differences in detection capabilities. For example, it is known that CPR has a much better

182  detection sensitivity compared with both DPR and ground radars in MRMS datasets (Wang et

183  al., 2018; Skofronick-Jackson et al., 2019).

184

185

186  **3 Results**

187    This section first uses a contingency table analysis to determine the snowfall success percentage

188    from the four datasets.  This analysis also explains why the snowfall success percentages differ

189    greatly across these datasets. Second, we analyze the snowfall (rainfall) success percentages

190    from six different phase determination schemes using ancillary parameters from four reanalysis

191    datasets.

192

193    **3.1 Snowfall success percentage**

194     As mentioned previously, there are two methods used in the DPR product to determine the

195    precipitation phase. The first method (hereafter referred to as DPR M1) separates the snowfall

196    from rainfall using temperature information at about 1.5 km above the ground (clutter-free

197    height), combined with bright-band information if it exists in the radar reflectivity profile. The

198    second method uses the KuPR and KaPR radar reflectivity difference at about 1.5 km above the

199    ground (hereafter referred to as DPR M2).

200

201    Figure 1 shows the snowfall success percentage from CPR, DPR, GPROF, and MRMS. It is

202    immediately clear that both DPR methods show rather poor performance with snowfall success

203    percentages of less than 50%, which means that more than half of the snowfall indicated by DPR

204    is rainfall or mixed precipiation at the surface, which is somewhat expected since both DPR

205    methods diagnose the precipitation phase at ~1.5 km above the ground, instead of at the surface.

206    The temperature can increase about 9 $^{\circ}$C from 1.5 km to the ground assuming a 6 ℃ temperature

207    lapse rate. In contrast, CPR demonstrates a much larger snowfall success percentage of about

208    96%. Over land, CPR uses a conservative classification scheme where pixels are classified as

209    snow if the maximum temperature in the column is less than 0 ℃, rain if the maximum

210    temperature in the column is greater than 2 ℃, and an undetermined "mixed precipitation" in all

211    other cases. This explains why CPR has a much better snowfall success percentage, compared

212    with DPR. The more accurate temperature information from the ECMWF analysis dataset used

213    in the CPR precipitation phase determination may also contribute to its better performance (more

214    details in the next section).

215

216    In the GPROF GMI retrievals, there are two precipitation rate variables (i.e., precipitation rate

217    and snowfall rate). First, we select snowfall rates greater than 0 to compute the snowfall success

218    percentage (hereafter referred to as GPROF M1). Second, we select pixels with a snowfall rate

219    greater than 0 and the difference between snowfall rate and precipitation rate is less than 0.01

220    mm/hr (hereafter referred to as GPROF M2). Figure 1 shows that GPROF M2 has a much higher

221    snow success percentage of 96% than that from GPROF M1 of 81%. The much larger success

222    percentage from GPROF M2 is directly determined by how GPROF calculates the snowfall

223    rates. Specifically, GPROF produces the snowfall rates by multiplying the precipitation rates

224    with the snowfall probability determined by the 2-m wet bulb temperature (Sims and Liu, 2015).

225    In the GPROF M2, we intentionally make these two variables close (i.e., the difference between

226    snowfall rate and precipitation rate less than 0.01 mm/hr). As expected, choosing different

227    threshold values (e.g., 0.1 or 0.001 mm/hr) will affect the snowfall success percentage values. In

228    fact, the snowfall success percentage is 99% (90%) with the difference being 0.001 mm/hr (0.1

229    mm/hr).  It is worth mentioning that evaluating the snowfall success percentage from GPROF is

230    actually equivalent to evaluating the snowfall probability computed from 2-m wet bulb

231    temperature based on Sims and Liu (2015).

232

233

234    For MRMS, the snowfall success percentage is 84%, while the vast majority of the other 16% of

235    data (i.e., surface indicates snowfall while MRMS judges them as rainfall) is misclassified as

236    "cold stratiform" rainfall. Using independent precipitation type reports from the citizen-

237    scientists, Chen et al., (2018) also noticed that MRMS tends to misidentify snowfall as rainfall.

238    The authors pointed out that the temperature threshold values or the uncertainties in the model-

239    output temperature information may be responsible for the misidentification. Further analysis

240    shows that the temperature information used in MRMS agrees very well with ground

241    observation. Therefore, we conclude that the most likely reason for the misclassification (i.e.,

242    snowfall as rainfall) in MRMS is caused by the "colder" threshold values.  Finally, it is worth

243    mentioning that the rainfall success percentages from all four datasets are greater than 94% with

244    no clear differences.

245

246    **3.2 Phase discrimination based on different methods and different reanalysis datasets**

247    In this section, we compute the snowfall and rainfall success percentages, using six phase

248    discrimination methods and four global reanalysis datasets. The first 5 approaches are from Sims

249    and Liu (2015), which uses 2-m air temperature (T2m), 2-m wet bulb temperature (Tw), 2-m wet

250    bulb temperature (Tw) & surface temperature (Ts), 2-m wet bulb temperature (Tw) &

251    temperature lapse rate ($\Gamma$), and 2-m wet bulb temperature (Tw) & temperature lapse rate ($\Gamma$) &

252    surface temperature (Ts), respectively.  It worth mentioning that the GPROF precipitation

253    product utilizes the second method (i.e., Tw) for the phase discrimination. These five methods, in

254    order, are referred to as M1, M2, ..., M5. The sixth method (referred to as M6) is based on

255    Haynes et al. (2009), which is the phase segregation method for CloudSat precipitation products.

256    As mentioned previously, CPR judged a pixel as the snowfall pixel over land when the

257    maximum temperature in the temperature profile less than 0º C (i.e., all temperature values are

258    less than 0º C in the profile). We do not include the MRMS phase discrimination method since

259    its threshold values are based on the temperature information from a regional model output.

260    Additionally, we only utilize the ground weather reports in 2016 since there are about 2.2 million

261    precipitating reports in total.

262

263    Table 1 shows the snowfall success percentage from these six different methods and four

264    different datasets. The results show that using 2-m wet-bulb temperature (M2) generates slightly

265    better phase discrimination performance than that using 2-m temperature (M1), regardless of the

266    reanalysis datasets. This is because that the wet-bulb temperature is closer to the hydrometeors'

267    temperature than the temperature itself since it combines the temperature and moisture

268    information (Sims and Liu, 2015). More importantly, we do not notice additional discriminant

269    capability being added when including more variables (e.g., lapse rate and surface temperature),

270    by comparing M2 with M3 to M5.  Further, the snowfall success percentage from M6 is slightly

271    lower than those from M1 to M5, regardless of the reanalysis datasets. For example, the success

272    percentage from M6 is 93% using ERA5 (Table 1, 7th row and 5th column), while it is 95% from

273    M1 (Table 1, 2nd row and 5th column). The reason why M6 generates slightly lower success

274    percentage is because M6 requires the maximum temperature in the profile being less than 0 ºC,

275    which is colder than the temperature threshold values used from M1 to M5 (e. g., 2-m air

276    temperature being 1.6 ºC in M1). Consequently, a small percentage of observed snowfall is

277    misclassified as mixed precipitation.

278

279     Table 2 shows that the rainfall success percentage based on MERRA2 is smaller than those from

280     other reanalysis datasets for M1 to M5, because MERRA2's temperature profile from about

281     200m to the surface is noticeably colder than those from the other three datasets, shown in Fig.

282     2a. The colder temperature from MERRA2 is further corroborated from the 2-m air temperature

283     plot in Fig. 2b, which demonstrates the 2-m air temperature difference between the reference and

284     those from each reanalysis datasets. Clearly, 2-m air temperature from MERRA2 is about 1 °C

285     colder than those from other datasets. Basically, the colder MERRA2 temperature misidentifies

286     the rainfall pixels as the snowfall pixels. The colder surface temperature from MERRA2 is also

287     reported by Draper et al., 2018.

288

289     Rainfall success percentage from M6 is about 92% using all four model outputs (Table 2). This

290     number (92%) is about 4% higher than those from M1 to M5 when using MERRA2, which is

291     because M6 requires the maximum temperature (Tmax) in the temperature profile greater than 2

292     °C. Even though MERRA2 is noticeably colder than observed, M6 achieves a slightly better

293     success percentage by using a slightly higher temperature threshold value than those from M1 to

294     M5 (e. g., 2-m air temperature being 1.6 ℃ in M1).  In contrast, M6 generates a slightly lower

295     success percentage comparing with M1 to M5 when using the other model outputs, which again

296     can be attributed to the 2 °C threshold value used in M6, which aims to exclude purely snowing

297     scenes while allowing a larger unknown "mixed" category. Basically, M6 classifies a small

298     percentage of rainfall as "mixed precipitation".

299

300     **4. Conclusions**

301     This study evaluated the precipitation phase discrimination performance in four widely used

302    remote-sensing precipitation datasets using global ground precipitation phase reports. It is found

303    that more than half of the snowfall indicated by GPM DPR is actually rainfall on the ground

304    because the phase determined by DPR is at about 1.5 km above the ground level. In other words,

305    more than half of the snowfall from DPR at the 1.5 km above the ground level melts to rain

306    drops at the surface level. In contrast, CloudSat CPR shows a much better snowfall classification

307    primarily because it considers the temperature profile information. In addition, the snowfall rate

308    in GPM GMI precipitation product is computed as the product of the snowfall probability and

309    the precipitation rate. Therefore, the snowfall discrimination accuracy is determined by the

310    snowfall probability accuracy calculated by Sims and Liu's method from the 2-m wet bulb

311    temperature (Sims and Liu, 2015). Further, MRMS misclassifies some snowfall as cold

312    stratiform rainfall. An analysis of four different reanalysis datasets shows that the MERRA2

313    temperature profile is noticeably colder close to the surface than those from the other three

314    global reanalysis datasets, which results in the rainfall being misclassified as snowfall. Finally, it

315    is found that using the 2-m wet-bulb temperature is adequate for snowfall determination

316    regardless of which reanalysis datasets is chosen.

317

318    The comparison analysis reveals that different remote sensing datasets use very different

319    approaches for precipitation phase determination, leading to significantly different accuracy

320    performance. Understanding this feature is critically important for correctly using these datasets

321    and future satellite precipitation dataset evaluations. This study also highlights the challenges

322    and obstacles in precipitation phase discrimination, an issue that must be addressed to generate

323    more accurate global snowfall climatology.

324

**References:**

Adhikari, A., Liu, C., & Kulie, M. S. (2018). Global distribution of snow precipitation features and their properties from 3 years of GPM observations. *J. Clim.*, 31 (10), 3731–3754.

Auer Jr, A. H. (1974). The rain versus snow threshold temperatures. Weatherwise, 27 (2), 67–67.

Behrangi, A., Yin, X., Rajagopal, S., Stampoulis, D., & Ye, H. (2018). On distinguishing snowfall from rainfall using near-surface atmospheric information: Comparative analysis, uncertainties and hydrologic importance. *Quart. J. Roy. Meteor. Soc.*, 144 , 89–102.

Bourgouin, P. (2000). A method to determine precipitation types. Wea. *Forecasting*, 15 (5), 583–592.

Chen, S., Gourley, J. J., Hong, Y., Cao, Q., Carr, N., Kirstetter, P.-E., . . . Flamig, Z. (2016). Using citizen science reports to evaluate estimates of surface precipitation type. *Bull. Amer. Meteor. Soc.*, 97 (2), 187–193.

Chen, S., Hong, Y., Kulie, M., Behrangi, A., Stepanian, P. M., Cao, Q., . . . Zhang, X. (2016). Comparison of snowfall estimates from the NASA CloudSat cloud profiling radar and NOAA/NSSL multi-radar multi-sensor system. *J. Hydrol.*, 541, 862–872.

Dai, A. (2001). Global precipitation and thunderstorm frequencies. Part I: Seasonal and interannual variations. *J. Clim.*, 14 (6), 1092–1111. Dai, A. (2008). Temperature and pressure dependence of the rain-snow phase transition over land and ocean. *Geophys. Res. Lett.*, 35 (12).

Ding, B., Yang, K., Qin, J., Wang, L., Chen, Y., & He, X. (2014). The dependence of precipitation types on surface elevation and meteorological conditions and its parameterization. *J. Hydrol.*, 513 , 154–163.

Draper, C.S., Reichle, R.H. and Koster, R.D., (2018). Assessment of MERRA-2 land surface energy flux estimates. *J. Clim.*, 31(2), 671-691.

371   Ebtehaj, A., & Kummerow, C. (2017). Microwave retrievals of terrestrial precipi- tation over

372   snow-covered surfaces: A lesson from the GPM satellite. *Geophys. Res. Lett.*, 44 (12), 6154–

373   6162.


374   Ferraro, R., Beauchamp, J., Cecil, D., &  Heymsfield, G.    (2015). A prototype hail detection

375   algorithm and hail climatology developed with the Advanced Microwave Sounding Unit

376   (AMSU). *Atmospheric Research*, 163, 24–35.


377   Gelaro, R., McCarty, W., Sua´rez, M. J., Todling, R., Molod, A., Takacs, L., . . . others (2017).

378   The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J.*

379   *Clim.*, 30 (14), 5419–5454.


380   Gergel, D. R., Nijssen, B., Abatzoglou, J. T., Lettenmaier, D. P., & Stumbaugh, M. R. (2017).

381   Effects of climate change on snowpack and fire potential in the western USA. *Climatic Change*,

382   141 (2), 287–299.


383   Hamada, A.,  & Takayabu,  Y. N. (2016). Improvements in Detection of Light Precipitation with

384   the Global Precipitation Measurement Dual-Frequency Precipitation Radar (GPM DPR). *J.*

385   *Atmos. Oceanic Technol.*, 33 (4), 653–667.


386   Haynes, J. M., L'Ecuyer, T. S., Stephens, G. L., Miller, S. D., Mitrescu, C., Wood, N. B., &

387   Tanelli, S. (2009). Rainfall retrieval over the ocean with spaceborne W-band radar. *J. Geophys.*

388   *Res.*, 114 (D8).


389   Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Hora´nyi, A., Mun˜oz-Sabater, J., . . . others

390   (2020). The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, 146 (730), 1999–2049.


391   Iguchi, T. (2020). Dual-Frequency Precipitation Radar (DPR) on the Global Precipitation

392   Measurement (GPM) Missions Core Observatory. In Satellite precipitation measurement (pp.

393   183–192). Springer.


394   Jennings, S., S., Winchell, B.,  Livneh, and P., Molotch, (2018). Spatial variation of the rain–

395    snow temperature threshold across the Northern Hemisphere. *Nat. comm.*, 9(1), 1-9.

396    Kanamitsu, I., Alpert, J.C., Campana, K.A., Caplan, P.M., Deaven, D.G., Iredell, M., Katz, B.,
397    Pan, H.L., Sela, J. and White, G.H., (1991). Recent changes implemented into the global forecast
398    system at NMC. *Wea. and Fore.*, 6(3), pp.425-435.

399    Kidd, C., Matsui, T., Chern, J., Mohr, K., Kummerow, C., & Randel, D.   (2016). Global
400    precipitation estimates from cross-track passive microwave observations using a physically
401    based retrieval scheme. *J. Hydrometeor.*, 17 (1), 383–400.

402    Kirstetter, P., Hong, Y., Gourley, J.J., Chen, S., Flamig, Z., Zhang, J., Schwaller, M., Petersen,
403    W. and Amitai, E., (2012). Toward a framework for systematic error modeling of spaceborne
404    precipitation radar with NOAA/NSSL ground radar–based National Mosaic QPE. *J.*
405    *Hydrometeor.*, 13(4), 1285-1300.

406    Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., . . . others (2015).  The
407    JRA-55 reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*. Ser. II
408    , 93 (1), 5–48.

409    Kulie, M. S., &  Bennartz, R. (2009). Utilizing spaceborne radars to retrieve dry snowfall. *J.*
410    *Appl. Meteor. Climatol.*, 48 (12), 2564–2580.

411    Kulie, M. S., Milani, L., Wood, N. B., Tushaus, S. A., Bennartz, R., & LEcuyer, T. S. (2016). A
412    Shallow Cumuliform Snowfall Census Using Spaceborne Radar. *J. Hydrometeor.*, 17 (4), 1261–
413    1279.

414    Kulie, M.S., Pettersen, C., Merrelli, A.J., Wagner, T.J., Wood, N.B., Dutter, M., Beachler, D.,
415    Kluber, T., Turner, R., Mateling, M. and Lenters, J., (2021). Snowfall in the Northern Great
416    Lakes: Lessons Learned from a Multi-Sensor Observatory. *Bull. Amer. Meteor. Soc.*, 1-61.

417    Kummerow, C. D., Randel, D. L., Kulie, M., Wang, N.-Y., Ferraro, R., Joseph Mun- chak, S., &
418    Petkovic, V. (2015). The evolution of the Goddard profiling algorithm to a fully parametric

419    scheme. *J. Atmos. Oceanic Technol.*, 32 (12), 2265–2280.

420    Le, M., Chandrasekar, V., & Biswas, S. (2017). An algorithm to identify surface snowfall from

421    GPM DPR observations. *IEEE Trans. Geosci. Remote Sens.*, 55 (7), 4059–4071.

422    Liu, G. (2008). Deriving snow cloud characteristics from CloudSat observations. *J. Geophys.*

423    *Res.*, 113 (D8).

424    Matsuo, T., Sasyo, Y., & Sato, Y. (1981). Relationship between types of precipitation on the

425    ground and surface meteorological elements. *J. Meteor. Soc. Japan*, 59 (4), 462–476.

426    Meng, H., Dong, J., Ferraro, R., Yan, B., Zhao, L., Kongoli, C., . . . Zavodsky, B. (2017). A

427    1DVAR-based snowfall rate retrieval algorithm for passive microwave radiometers. *J. Geophys.*

428    *Res.*, 122 , 6520-6540.

429    Milani, L., Kulie, S., Casella, D., Dietrich, S., L'Ecuyer, T.S., Panegrossi, G., Porcù, F., Sanò, P.

430    and Wood, B., (2018). CloudSat snowfall estimates over Antarctica and the Southern Ocean: An

431    assessment of independent retrieval methodologies and multi-year snowfall analysis. *Atmos.*

432    *Res.*, 213, 121-135.

433    Milani, L., Kulie, M.S., Casella, D., Kirstetter, P.E., Panegrossi, G., Petkovic, V., Ringerud, S.E.,

434    Rysman, J.F., Sanò, P., Wang, N.Y. and You, Y., (2021). Extreme Lake-Effect Snow from a

435    GPM Microwave Imager Perspective: Observational Analysis and Precipitation Retrieval

436    Evaluation. *J. Atmos. Oceanic Technol.*, 38(2), 293-311.

437    Mroz, K., Battaglia, A., Lang, T.J., Cecil, D.J., Tanelli, S. and Tridon, F., (2017). Hail-detection

438    algorithm for the GPM Core Observatory satellite sensors. *J. Appl. Meteorol. Clim.*, 56(7),

439    1939-1957.

440    United States Army Corps of Engineers (1956). Summary report of the snow investigation Snow

441    hydrology. North Pacific Division report,437 pp., Washington, D. C..

442 OGorman, P. A. (2014). Contrasting responses of mean and extreme snowfall to climate change.
443 *Nat.*, 512 (7515), 416–418.

444 Panegrossi, G., Rysman, J.-F., Casella, D., Marra, A. C., San`o, P., & Kulie, M. S. (2017).
445 CloudSat-based assessment of GPM Microwave Imager snowfall observation capabilities.
446 *Remote  Sens.*, 9 (12), 1263.

447 Petty, G. W.  (1995).  Frequencies and characteristics of global oceanic precipitation from
448 shipboard present-weather reports. *Bull. Amer. Meteor. Soc.*, 76 (9), 1593–1616.

449 Sheffield, J., Ferguson, C. R., Troy, T. J., Wood, E. F., & McCabe, M. F. (2009). Closing the
450 terrestrial water budget from satellite remote sensing.  *Geophys. Res. Lett.*, 36 (7).

451 Sims, M., and Liu, G. (2015). A parameterization of the probability of snow–rain transition. *J.*
452 *Hydrometeor.*, 16 (4), 1466–1477.

453 Skofronick-Jackson, G., Kulie, M., Milani, L., Munchak, S. J., Wood, N. B., & Levizzani, V.
454 (2019). Satellite estimation of falling snow: A global precipitation measurement (GPM) core
455 observatory perspective. *J. Appl. Meteor. Climatol.*, 58 (7), 1429–1448.

456 Smith, A., Lott, N., & Vose, R. (2011). The integrated surface database: Recent de- velopments
457 and partnerships. *Bull. Amer. Meteor. Soc.*, 92 (6), 704–708.

458 Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., . . . others  (2002).
459 The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds
460 and precipitation. *Bull. Amer. Meteor. Soc.*, 83 (12), 1771–1790.

461 Wagner, A. J. (1957). Mean temperature from 1000 mb to 500 mb as a predictor of precipitation
462 type. *Bull. Amer. Meteor. Soc.*, 38 (10), 584–590.

463 Wang, Y., Liu, G., Seo, E.-K., & Fu, Y. (2013). Liquid water in snowing clouds: Im- plications
464 for satellite remote sensing of snowfall. *Atmos. Res.*, 131, 60–72.

465 Wang, Y., You, Y., & Kulie, M. (2018).Global Virga Precipitation Distribution Derived

466 From Three Spaceborne Radars and Its Contribution to the False Radiometer Precipitation

467 Detection. *Geophys. Res. Lett.*, 45 (9), 4446–4455.

468 You, Y., Wang, N.-Y., & Ferraro, R. (2015). A Prototype Precipitation Retrieval Algorithm Over

469 Land Using Passive Microwave Observations Stratified by Sur- face Condition and Precipitation

470 Vertical Structure. *J. Geophys. Res.*, 120, 5295–5315.

471 You, Y., Wang, N.-Y., Ferraro, R., & Meyers, P. (2016). A Prototype Precipitation Retrieval

472 Algorithm over Land for ATMS. *J. Hydrometeor.*, 17 (5), 1601–1621.

473 You, Y., Wang, N.-Y., Ferraro, R., & Rudlosky, S. (2017). Quantifying the Snowfall Detection

474 Performance of the GPM Microwave Imager Channels over Land. *J. Hydrometeor.*, 18 (3), 729–

475 751.

476 Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., . . . others (2016). Multi-Radar

477 Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull.*

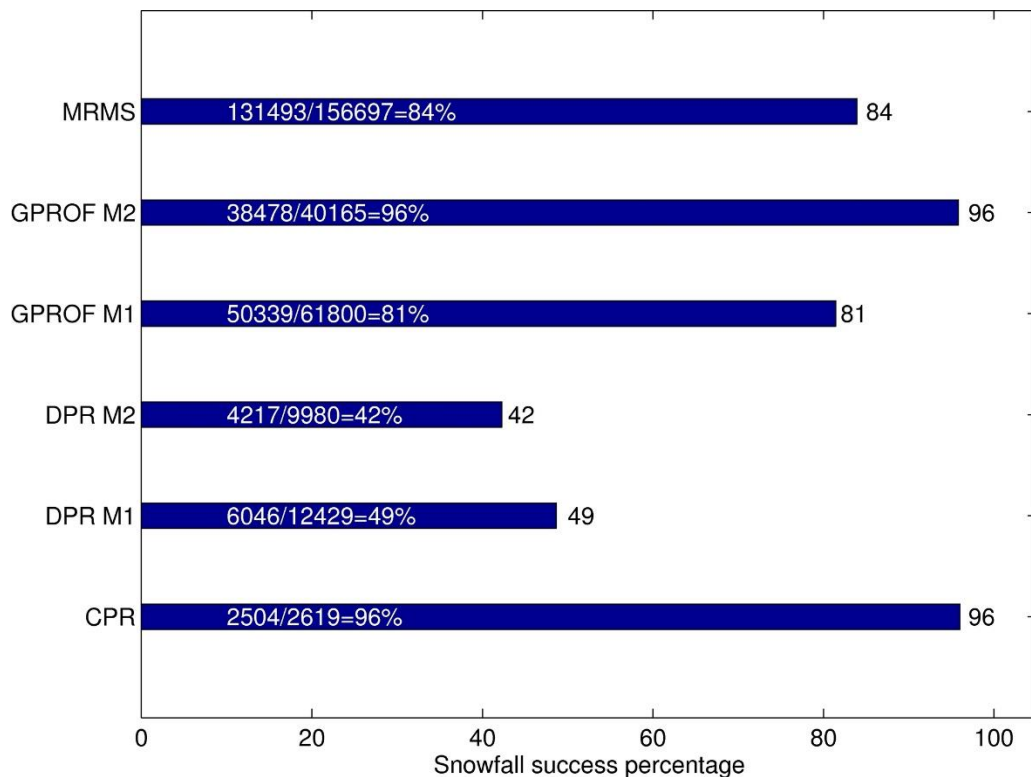478 *Amer. Meteor. Soc.*, 97 (4), 621–638.

479 Table 1. Snowfall success percentage (%) from six methods (M1 to M6) and four reanalysis

480 datasets (MERRA2, JRA55, ERA5, and GFS).  The first 5 approaches are from Sims and Liu

481 (2015), which use 2-m air temperature (T2m), 2-m wet bulb temperature (Tw), 2-m wet bulb

482 temperature (Tw) & surface temperature (Ts), 2-m wet bulb temperature (Tw) & temperature

483 lapse rate ($\Gamma$), and 2-m wet bulb temperature (Tw) & temperature lapse rate ($\Gamma$) & surface

484 temperature (Ts), respectively. The sixth approach is from Haynes et al. (2009), which requires

485 the maximum temperature (Tmax) in the temperature profile being less than 0 °C.

486

| Method | Variables | MERRA2 | JRA55 | ERA5 | GFS |
|--------|-----------|--------|-------|------|-----|
| M1 | T2m | 94 | 93 | 95 | 95 |
| M2 | Tw | 97 | 95 | 96 | 96 |
| M3 | Tw, Ts | 97 | 94 | 96 | 95 |
| M4 | Tw, $\Gamma$ | 97 | 95 | 96 | 96 |
| M5 | Tw, $\Gamma$, Ts | 97 | 95 | 96 | 96 |
| M6 | Tmax | 91 | 90 | 93 | 91 |

487

488

489 Table 2. Rainfall success percentage (%) from six methods (M1 to M6) and four reanalysis

490 datasets (MERRA2, JRA55, ERA5, and GFS).  The first 5 approaches are from Sims and Liu

491 (2015), which use 2m air temperature (T2m), 2-m wet bulb temperature (Tw), 2-m wet bulb

492 temperature (Tw) & surface temperature (Ts), 2-m wet bulb temperature (Tw) & temperature

493 lapse rate ($\Gamma$), and 2-m wet bulb temperature (Tw) & temperature lapse rate ($\Gamma$) & surface

494 temperature (Ts), respectively. The sixth approach is from Haynes et al. (2009), which requires

495 the maximum temperature (Tmax) in the temperature profile greater than 2 °C.

496

| Method | Variables | MERRA2 | JRA55 | ERA5 | GFS |
|--------|-----------|--------|-------|------|-----|
| M1 | T2m | 89 | 95 | 96 | 96 |
| M2 | Tw | 88 | 94 | 96 | 96 |
| M3 | Tw, Ts | 88 | 95 | 97 | 96 |
| M4 | Tw, $\Gamma$ | 88 | 96 | 97 | 97 |
| M5 | Tw, $\Gamma$, Ts | 88 | 97 | 97 | 97 |
| M6 | Tmax | 92 | 91 | 93 | 92 |

497
498
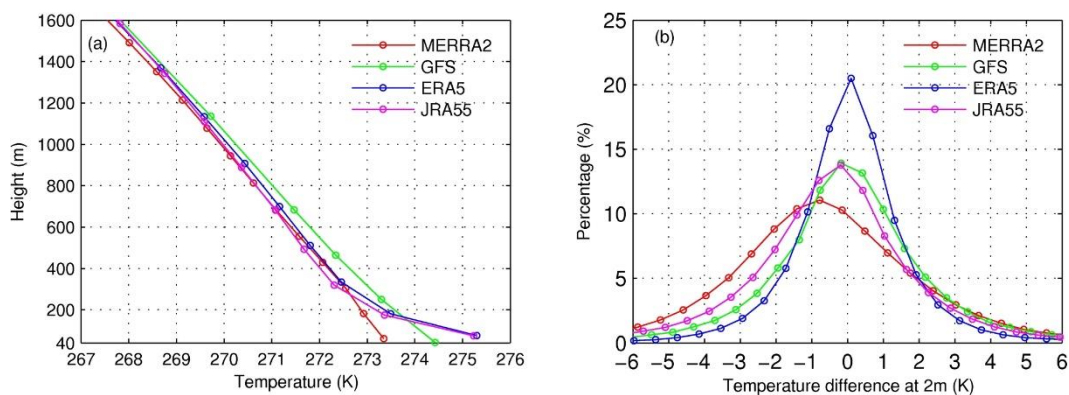
499

500



Figure 1. The snowfall success percentage from four precipitation remote sensing datasets,

including CloudSat Cloud Profiling Radar (CPR), GPM Dual Frequency Precipitation Radar

(DPR), GPM GMI precipitation retrieval results from the Goddard Profiling Algorithm (GPROF)

algorithm, and the Multi-Radar Multi-Sensor (MRMS). The denominator on each bar is the

snowfall number from the ground weather report, while the numerator is the number from each

remote sensing dataset. DPR M1 and DPR M2 represent two methods used in DPR for the

precipitation phase discrimination. GPROF M1 and DPR M2 represent two methods used in

GPROF retrieval for the precipitation phase discrimination (see corresponding text for more

details). MRMS data are only available over the Continental United States (CONUS).

511

512



Figure 2.  (a) Temperature profiles below 1600 meters from four reanalysis datasets, including

MERRA2, GFS, ERA5, and JRA55. (b) The histograms of the 2-m air temperature differences

between surface observations and each reanalysis dataset. All data are from January 2016 to

December 2016.